Wavelet Based Semi-blind Channel Estimation For Multiband OFDM

Sajad Sadough[†*], Mahieddine M. Ichir[*], Emmanuel Jaffrot[†] and Pierre Duhamel[*]
[†] UEI, ENSTA, 32 boulevard Victor, 75015 Paris, France
[*] Laboratoire des Signaux et Systèmes, CNRS-Supélec,
Plateau de Moulon, 91190 Gif–sur–Yvette, France
Email: sajad.sadough@ensta.fr

arXiv:0708.1414v1 [cs.NI] 10 Aug 2007

***Abstract*— This paper introduces an *expectation-maximization* (EM) algorithm within a wavelet domain Bayesian framework for semi-blind channel estimation of multiband OFDM based UWB communications. A prior distribution is chosen for the wavelet coefficients of the unknown channel impulse response in order to model a sparseness property of the wavelet representation. This prior yields, in *maximum a posteriori* estimation, a thresholding rule within the EM algorithm. We particularly focus on reducing the number of estimated parameters by iteratively discarding "unsignificant" wavelet coefficients from the estimation process. Simulation results using UWB channels issued from both models and measurements show that under sparsity conditions, the proposed algorithm outperforms pilot based channel estimation in terms of mean square error and bit error rate and enhances the estimation accuracy with less computational complexity than traditional semi-blind methods.**

I. Introduction

A UWB radio signal is defined as any signal whose bandwidth is larger than 20% of its center frequency or greater than 500 MHz [1]. In recent years, UWB system design has experienced a shift from the traditional "single-band" radio that occupies the whole 7.5 GHz allocated spectrum to a "multiband" design approach [2]. That consists in dividing the available UWB spectrum into several subbands, each one occupying approximately 500 MHz.

Multiband Orthogonal Frequency Division Multiplexing (MB-OFDM) [3] is a strong candidate for multiband UWB which enables high data rate UWB transmission to inherit all the strength of OFDM that has already been shown for wireless communications (ADSL, DVB, 802.11a, 802.16.a, etc.). This approach uses a conventional coded OFDM system [4] together with bit interleaved coded modulation (BICM) and frequency hopping over different subbands to improve diversity and to enable multiple access.

Basic receivers proposed for MB-OFDM [3], estimate the channel by using pilots (known training symbols) transmitted at the beginning of the information frame, implicitly assuming a time invariant channel within a single frame. Thus, for an accurate channel acquisition, one must send several pilot patterns resulting in a significant loss in spectral efficiency.

Recent works [5], [6] have reported promising results on the combination of channel estimation and data decoding process by using the Expectation-Maximization (EM) algorithm [7] . Though the latter scheme outperforms pilot based receivers, it has a higher complexity that may be of a critical concern for its practical implementations. This complexity is mainly dominated by the number of estimated parameters for channel updating and the decoding algorithm within each iteration.

In this work, we consider a semi-blind joint channel estimation and data detection scheme based on the EM algorithm, with the objective of minimizing the number of estimated parameters and enhancing the estimation accuracy. This is achieved by expressing the unknown channel impulse response (CIR) in terms of its discrete wavelet series, which has been shown to provide a *parsimonious* representation [8], [9]. Thus, we choose a particular prior distribution for the channel wavelet coefficients that renders the maximum a posteriori (MAP) channel estimation equivalent to a hard thresholding rule at each iteration of the EM algorithm. The latter is then exploited to reduce the estimator computational load by discarding "unsignificant" wavelet coefficients from the estimation process. Moreover, since the probability of encoded bits are involved in the EM computation, we naturally combine the iterative process of channel estimation with the decoding operation of encoded data.

This paper is organized as follows. Section II introduces MB-OFDM and its wavelet domain channel estimation observation model. In section III, we first describe a MAP version of the EM algorithm for channel estimation and then show how the number of estimated parameters can be reduced through the EM iterations. The combination of the channel estimation part with the decoding operation and implementation issues are also discussed. Section V illustrates, via simulations, the performance of the proposed receiver over a realistic UWB channel environment and section VI concludes the paper.

Notational conventions are as follows: $\mathcal{D}_{\mathbf{x}}$ is a diagonal matrix with diagonal elements $\mathbf{x} = [x_1, \ldots, x_N]^T$, $\mathbb{E}_{\mathbf{x}}[.]$ refers to expectation with respect to $\mathbf{x}$, $\mathbf{I}_N$ denotes an $(N \times N)$ identity matrix; $\|.\|$, $(.)^*$, $(.)^T$ and $(.)^{\mathcal{H}}$ denote Frobenious norm, matrix or vector conjugation, transpose and Hermitian transpose, respectively.

II. System model and wavelet domain problem formulation

MB-OFDM system divides the spectrum between 3.1 to 10.6 GHz into several non-overlapping subbands each one occupying 528 MHz of bandwidth [3]. The transmitter architecture for the MB-OFDM system is very similar to that of a

conventional wireless OFDM system. The main difference is that MB-OFDM system uses a time-frequency code (TFC) to select the center frequency of different subbands which is used not only to provide frequency diversity but also to distinguish between multiple users (see figures 1 and 2). Here, we consider MB-OFDM in its basic mode *ie.* employing the three first subbands.

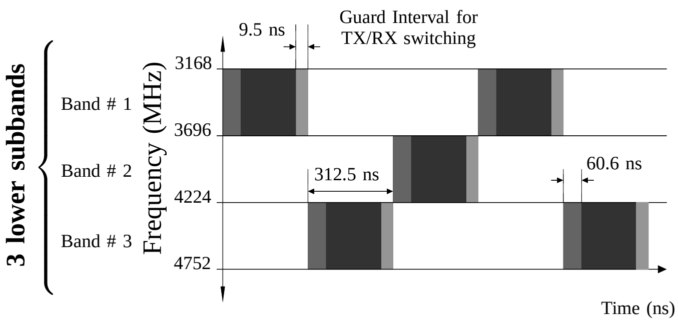

Fig. 1. Example of time-frequency coding for the multiband OFDM system: TFC={1, 3, 2, 1, 3, 2, ...}.

We consider the multiband OFDM transmission of figure 2 using N data subcarriers. At the receiver, assuming a cyclic prefix (CP) longer than the channel maximum delay spread and perfect synchronization, OFDM converts a frequency selective channel into N parallel flat fading subchannels [4] for each subband as

$$\mathbf{y}_{i,n} = \mathcal{D}_{\mathbf{s}_{i,n}}\,\mathbf{h}_{i,n} + \mathbf{z}_{i,n} \quad i \in \{1,2,3\}, \quad n = 1, \ldots, \mathrm{N_{sym}} \tag{1}$$

where $(1 \times N)$ vectors $\mathbf{y}_{i,n}$, $\mathbf{s}_{i,n}$ and $\mathbf{h}_{i,n}$ denote received and transmitted symbols, and the channel frequency response respectively; the noise block $\mathbf{z}_{i,n}$ is assumed to be a zero mean white complex Gaussian noise with distribution $\mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I}_N)$; i is the subband index and n refers to the OFDM symbol index inside the frame. The observation model corresponding to all three subbands can be written in frequency domain as

$$\mathbf{Y}_m = \mathcal{D}_{\mathbf{s}_m}\,\mathbf{H}_m + \mathbf{Z}_m \quad m = 1, \ldots, \mathrm{M_{sym}} \tag{2}$$

where $\mathbf{Y}_m = [\mathbf{y}_{1,n}, \mathbf{y}_{2,n}, \mathbf{y}_{3,n}]^T$, $\mathbf{S}_m = [\mathbf{s}_{1,n}, \mathbf{s}_{2,n}, \mathbf{s}_{3,n}]^T$, $\mathbf{H}_m = [\mathbf{h}_{1,n}, \mathbf{h}_{2,n}, \mathbf{h}_{3,n}]^T$ and $\mathbf{Z}_m = [\mathbf{z}_{1,n}, \mathbf{z}_{2,n}, \mathbf{z}_{3,n}]^T$ are $(M \times 1)$ vectors, with $M = 3N$ and $\mathrm{M_{sym}} = \mathrm{N_{sym}}/3$. In the remainder, unless otherwise mentioned, we will not write the time index m for notational convenience.

In order to take advantage of the wavelet based estimation, the channel impulse response is expressed in terms of its orthogonal discrete wavelet coefficients. Let $\mathbf{F}_{M,L}$ be the truncated fast Fourier transform (FFT) matrix constructed from the $(M \times M)$ FFT matrix by keeping the first L columns where L is the length of the CIR over a group of three subbands. We define $\mathbf{W}$ as the $(L \times L)$ orthogonal discrete wavelet transform (ODWT) matrix. The unknown channel can be expressed as $\mathbf{H} = \mathbf{F}_{M,L}\mathbf{W}^{\mathcal{H}}\mathbf{g}$, where $\mathbf{g}$ is the $(L \times 1)$ vector of the CIR wavelet coefficients. The Observation model 2 is rewritten as

$$\mathbf{Y} = \mathcal{D}_{\mathbf{s}}\,\mathbf{T}\,\mathbf{g} + \mathbf{Z} \tag{3}$$

where $\mathbf{T} = \mathbf{F}_{M,L}\mathbf{W}^{\mathcal{H}}$.

Although at the transmitter, the channel is practically used by slices of 528 MHz bandwidth that corresponds to one of the subbands, at the receiver side we gather three received OFDM symbols for estimating the wavelet coefficients of the CIR, taken over all of the subbands (1.584 GHz bandwidth). This is motivated by the fact that estimating the channel over a wider bandwidth leads to a sparser wavelet representation. Besides, this approach simplifies the receiver architecture since there is no need to change the central frequency for down converting different subbands.

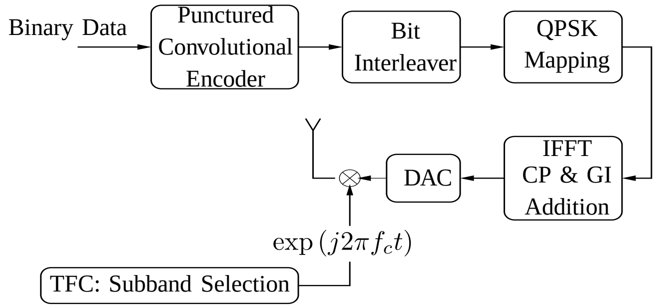

Fig. 2. TX architecture of the multiband OFDM system.

III. The EM-MAP algorithm for wavelet domain channel estimation

The EM algorithm proposed in this section is able to integrate the advantages of wavelet based estimation via the prior choosen for channel wavelet coefficients. Next, we see how the MAP estimator leads to a thresholding procedure which is used for reducing the number of estimated coefficients at each iteration of the EM algorithm.

A. An equivalent model and the EM principle

Our first step consists in decomposing the AWGN in (3) into the sum of two different Gaussian noise terms as

$$\mathbf{Z} = \mathcal{D}_{\mathbf{s}}\mathbf{Z}_1 + \mathbf{Z}_2 \tag{4}$$

where $\mathbf{Z}_1$ and $\mathbf{Z}_2$ are $(M \times 1)$ independent Gaussian noise vectors such that $p(\mathbf{Z}_1) = \mathcal{CN}(\mathbf{0}, \alpha^2\mathbf{I}_M)$ and $p(\mathbf{Z}_2) = \mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I}_M - \alpha^2\mathcal{D}_{\mathbf{s}}\mathcal{D}_{\mathbf{s}}^{\mathcal{H}})$. Since we are using normalized QPSK symbols, $\mathcal{D}_{\mathbf{s}}\mathcal{D}_{\mathbf{s}}^{\mathcal{H}} = \mathbf{I}_M$ and the covariance matrix of $\mathbf{Z}_2$ reduces to $\mathbf{\Sigma}_2 = (\sigma^2 - \alpha^2)\mathbf{I}_M$. We define the positive design parameter $\rho \triangleq \alpha^2/\sigma^2$, $(0 < \rho \leq 1)$ and notice that setting $\rho = 1$ leads to $\mathbf{Z}_2 = 0$ which is equivalent to working with the initial model (3). However, for $0 < \rho < 1$, the above noise decomposition allows the introduction of a hidden channel vector $\widetilde{\mathbf{H}}$ defined as

$$\begin{cases} \widetilde{\mathbf{H}} & = \mathbf{T}\,\mathbf{g} + \mathbf{Z}_1 \\ \mathbf{Y} & = \mathcal{D}_{\mathbf{s}}\,\widetilde{\mathbf{H}} + \mathbf{Z}_2. \end{cases} \tag{5}$$

The hidden vector $\widetilde{\mathbf{H}}$ provides a direct relation between true and estimated wavelet coefficients corrupted by an additive white Gaussian noise, allowing the two-stage observation model (5) which is equivalent to (3). However, the difference with a standard denoising problem is that $\mathbf{S}$ and $\widetilde{\mathbf{H}}$ are unknown. Hence, the observation model has missing datas and hidden variables and the MAP solution of $\mathbf{g}$ has no closed form. In such situations, the EM algorithm [7] is often used

to maximize the expectation of the posterior distribution over all possible missing and hidden variables.

Let $\mathbf{X} = \{\mathbf{Y}, \mathbf{S}, \widetilde{\mathbf{H}}\}$ be the *complete data set* in the EM algorithm terminology. Note that the observation set $\mathbf{Y}$ determines only a subset of the space $\mathscr{X}$ of which $\mathbf{X}$ is an outcome. We search $\mathbf{g}$ that maximizes $\log p(\mathbf{g}|\mathbf{X})$. After initialization by a short pilot sequence at the beginning of the frame, the EM algorithm alternates between the following two steps (until some stopping criterion) to produce a sequence of estimates $\{\mathbf{g}^{(t)},\, t = 0, 1, \ldots, t_{\max}\}$.

- **Expectation Step** (E-step): The conditional expectation of the complete log-likelihood given the observed vector and the current estimate $\mathbf{g}^{(t)}$ is calculated. This quantity is called the *auxiliary* or Q-function

$$Q\big(\mathbf{g}, \mathbf{g}^{(t)}\big) = \mathbb{E}_{\mathbf{S},\widetilde{\mathbf{H}}}\Big[\log p(\mathbf{Y}, \mathbf{S}, \widetilde{\mathbf{H}}|\mathbf{g})\Big|\mathbf{y}, \mathbf{g}^{(t)}\Big] \tag{6}$$

- **Maximization Step** (M-step): The estimated parameter is updated according to

$$\mathbf{g}^{(t+1)} = \arg\max_{\mathbf{g}} \Big\{Q\big(\mathbf{g}, \mathbf{g}^{(t)}\big) + \log \pi(\mathbf{g})\Big\} \tag{7}$$

where $\pi(\mathbf{g})$ is a *prior* distribution for the wavelet coefficients. Next, we derive the specific formulas of each step, according to (5).

B. E-step: Computation of the Q-function

The complete likelihood is

$$p(\mathbf{Y}, \mathbf{S}, \widetilde{\mathbf{H}}|\mathbf{g}) = p(\mathbf{Y}|\mathbf{S}, \widetilde{\mathbf{H}}, \mathbf{g})\, p(\mathbf{S}|\widetilde{\mathbf{H}}, \mathbf{g})\, p(\widetilde{\mathbf{H}}|\mathbf{g}).$$

According to (5), conditioned on $\widetilde{\mathbf{H}}$, $\mathbf{Y}$ is independent of $\mathbf{g}$. Furthermore, $\mathbf{S}$ which results from coding and interleaving of bit sequence is independent of $\widetilde{\mathbf{H}}$ and $\mathbf{g}$. Since $\mathbf{Z}_1$ is a complex white Gaussian noise, the complete log-likelihood can be simplified to

$$\begin{aligned}\log p(\mathbf{Y}, \mathbf{S}, \widetilde{\mathbf{H}}|\mathbf{g}) &= \log\big[p(\mathbf{Y}|\mathbf{S}, \widetilde{\mathbf{H}})\; p(\mathbf{S})\; p(\widetilde{\mathbf{H}}|\mathbf{g})\big] \\ &= \log p(\widetilde{\mathbf{H}}|\mathbf{g}) \;+\; \text{cst.} \\ &= -\frac{\mathbf{g}^{\mathcal{H}}\mathbf{T}^{\mathcal{H}}\mathbf{T}\mathbf{g} - 2\mathbf{g}^{\mathcal{H}}\mathbf{T}^{\mathcal{H}}\widetilde{\mathbf{H}}}{\alpha^2} \;+\; \text{cst.}\end{aligned} \tag{8}$$

where cst. are different constant terms that do not depend on $\mathbf{g}$. According to (6) we have

$$\begin{aligned}Q\big(\mathbf{g}, \mathbf{g}^{(t)}\big) &= \mathbb{E}_{\mathbf{S},\widetilde{\mathbf{H}}}\Big[-\frac{\mathbf{g}^{\mathcal{H}}\mathbf{T}^{\mathcal{H}}\mathbf{T}\mathbf{g} - 2\mathbf{g}^{\mathcal{H}}\mathbf{T}^{\mathcal{H}}\widetilde{\mathbf{H}}}{\alpha^2} + \text{cst.}\Big|\mathbf{Y}, \mathbf{g}^{(t)}\Big] \\ &= -\frac{\|\,\langle\widetilde{\mathbf{H}}^{(t)}\rangle - \mathbf{T}\mathbf{g}\,\|^2}{\alpha^2} + \text{cst.}\end{aligned} \tag{9}$$

where $\langle\widetilde{\mathbf{H}}^{(t)}\rangle \triangleq \mathbb{E}_{\mathbf{S},\widetilde{\mathbf{H}}}[\widetilde{\mathbf{H}}|\mathbf{Y}, \mathbf{g}^{(t)}]$.

From (9), it is obvious that the E-step involves only the computation of $\langle\widetilde{\mathbf{H}}^{(t)}\rangle$, we have

$$\langle\widetilde{\mathbf{H}}^{(t)}\rangle = \sum_{\mathbf{S}\in\mathscr{C}}\bigg(\int_{\widetilde{\mathbf{H}}\in\mathscr{H}}\widetilde{\mathbf{H}}\quad p(\widetilde{\mathbf{H}}|\mathbf{Y}, \mathbf{g}^{(t)})\,\mathrm{d}\widetilde{\mathbf{H}}\bigg)p(\mathbf{S}|\mathbf{Y}, \mathbf{g}^{(t)}) \tag{10}$$

where the last equation results from the independence between $\mathbf{S}$ and $\widetilde{\mathbf{H}}$ belonging respectively to the sets $\mathscr{C}$ and $\mathscr{H}$ which contain all of their possible values.

In order to evaluate $\langle\widetilde{\mathbf{H}}^{(t)}\rangle$, we first have to evaluate the conditional mean $\boldsymbol{\mu}_{\widetilde{\mathbf{H}}}^{(t)}$ of $\widetilde{\mathbf{H}}$ as

$$\boldsymbol{\mu}_{\widetilde{\mathbf{H}}}^{(t)} = \int_{\widetilde{\mathbf{H}}\in\mathscr{H}}\widetilde{\mathbf{H}}\quad p(\widetilde{\mathbf{H}}|\mathbf{Y}, \mathbf{g}^{(t)})\,\mathrm{d}\widetilde{\mathbf{H}} \tag{11}$$

Since both $p(\mathbf{Y}|\widetilde{\mathbf{H}})$ and $p(\widetilde{\mathbf{H}}|\mathbf{g}^{(t)})$ are Gaussian densities, $p(\widetilde{\mathbf{H}}|\mathbf{Y}, \mathbf{g}^{(t)}) \;\propto\; p(\mathbf{Y}|\widetilde{\mathbf{H}})\;\; p(\widetilde{\mathbf{H}}|\mathbf{g}^{(t)})$ is also Gaussian. By standard manipulation of Gaussian densities, we obtain

$$\boldsymbol{\mu}_h^{(t)} = \mathbf{T}\mathbf{g}^{(t)} + \rho\,\mathcal{D}_{\mathbf{s}}^{\mathcal{H}}\Big(\mathbf{Y} - \mathcal{D}_{\mathbf{s}}\mathbf{T}\mathbf{g}^{(t)}\Big). \tag{12}$$

By using (12) in (10) and after some simplifications we get

$$\langle\widetilde{\mathbf{H}}^{(t)}\rangle = (1-\rho)\,\mathbf{T}\mathbf{g}^{(t)} + \rho\,\overline{\mathcal{D}}_{\mathbf{s}}^{\mathcal{H}}\mathbf{Y} \tag{13}$$

where $\overline{\mathcal{D}}_{\mathbf{s}} = \sum_{\mathbf{s}\in\mathscr{C}}\mathcal{D}_{\mathbf{s}}\,p(\mathbf{S}|\mathbf{Y}, \mathbf{g}^{(t)})$.

The E-step is then completed by inserting $\langle\widetilde{\mathbf{H}}^{(t)}\rangle$ into $Q(\mathbf{g}, \mathbf{g}^{(t)})$, equation (9).

C. M-step: Wavelet Based MAP Estimation

In this step the estimate of the parameter $\mathbf{g}$ is updated as given in (7) where $Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(t)})$ is given by (9)

$$\mathbf{g}^{(t+1)} = \arg\max_{\mathbf{g}}\left\{-\frac{\|\,\langle\widetilde{\mathbf{H}}^{(t)}\rangle - \mathbf{T}\mathbf{g}\,\|^2}{\alpha^2} + \log\pi(\mathbf{g})\right\}. \tag{14}$$

Due to the orthonormality of both Fourier and wavelet transforms, $\mathbf{T}^{\mathcal{H}}\mathbf{T} = \mathbf{I}_L$ and we can replace $\|\,\langle\widetilde{\mathbf{H}}^{(t)}\rangle - \mathbf{T}\mathbf{g}\,\|^2$ by $\|\,\widetilde{\mathbf{g}}^{(t)} - \mathbf{g}\,\|^2$, where

$$\begin{aligned}\widetilde{\mathbf{g}}^{(t)} &= \mathbf{T}^{\mathcal{H}}\langle\widetilde{\mathbf{H}}^{(t)}\rangle \\ &= (1-\rho)\,\mathbf{g}^{(t)} + \rho\,(\overline{\mathcal{D}}_{\mathbf{s}}\mathbf{T})^{\mathcal{H}}\mathbf{Y}\end{aligned} \tag{15}$$

The M-step can be written as

$$\mathbf{g}^{(t+1)} = \arg\max_{\mathbf{g}}\left\{-\frac{\|\,\widetilde{\mathbf{g}}^{(t)} - \mathbf{g}\,\|^2}{\alpha^2} + \log\pi(\mathbf{g})\right\}. \tag{16}$$

Actually $\mathbf{g}^{(t+1)}$ in (14) is no more than the MAP estimate of $\mathbf{g}$ from the observation model

$$\widetilde{\mathbf{g}}^{(t)} = \mathbf{g} + \mathbf{Z}_1' \tag{17}$$

where $\mathbf{Z}_1' = \mathbf{T}^{\mathcal{H}}\mathbf{Z}_1 \sim \mathcal{CN}(\mathbf{0}, \alpha^2\mathbf{I}_L)$. From the Bayes theorem, the posterior distribution of $\mathbf{g}$ is given by

$$p\left(\mathbf{g}|\widetilde{\mathbf{g}}^{(t)}\right) \;\propto\; p\left(\widetilde{\mathbf{g}}^{(t)}|\mathbf{g}\right)\pi\left(\mathbf{g}\right) \tag{18}$$

where $p(\widetilde{\mathbf{g}}^{(t)}|\mathbf{g})$ is the Gaussian likelihood, $\widetilde{\mathbf{g}} \sim \mathcal{CN}(\mathbf{g}, \alpha^2\mathbf{I}_L)$. In this approach, we adopt the Bernoulli-Gaussian prior distribution $\pi\left(\mathbf{g}\right)$ for the wavelet coefficients $\mathbf{g}$ of the unknown CIR described by

$$\pi\left(g_j\right) = \lambda\;\delta(g_j) + (1-\lambda)\;\mathcal{CN}_{g_j}\left(0, \tau^2\right) \tag{19}$$

for $j = 1, \ldots, L$, which allows us to model a sparseness property of UWB channels in wavelet domain. This amounts considering that the wavelet coefficients have a probability λ to be zero and a probability $1-\lambda$ to be distributed as $\mathcal{CN}(0, \tau^2)$.

In order to deal with that particular model, we introduce an additional state variable (or indicator) $\beta_j \in \{0,1\}$ such that we can express this prior conditionally as

$$\begin{cases} (g_j|\beta_j = 0) \quad \sim \quad \delta(g_j) & \text{with probability } \lambda, \\ (g_j|\beta_j = 1) \quad \sim \quad \mathcal{CN}_{g_j}\left(0, \tau^2\right) & \text{with probability } 1-\lambda. \end{cases} \tag{20}$$

This prior model, conditionally on that state variable, leads to a Gaussian posterior for g_j which makes the estimation explicit; from the direct observation model $\widetilde{g}_j^{(t)} = g_j + Z'_{1,j}$, we can express these posterior probabilities of β_j as

$$\begin{aligned} p\left(\beta_j = 0|\widetilde{g}_j^{(t)}\right) &= \lambda\, \mathcal{N}\left(0, \alpha^2\right)/c \\ p\left(\beta_j = 1|\widetilde{g}_j^{(t)}\right) &= (1-\lambda)\, \mathcal{N}\left(0, \alpha^2+\tau^2\right)/c \end{aligned} \tag{21}$$

where the constant $c = \lambda \mathcal{N}\left(0,\alpha^2\right) + (1-\lambda)\,\mathcal{N}\left(0,\alpha^2+\tau^2\right)$. From this set of equations, we easily notice that the indicator variable β_j allows us to discriminate between the noise coefficients (for $\beta_j = 0$) and the effective channel wavelet coefficients (for $\beta_j = 1$), eventually corrupted by noise. The indicator variables β_j are estimated, in the MAP sense, by

$$\beta_j^{(t+1)} = \begin{cases} 0, & \text{if } \ p\left(\beta_j = 0|\widetilde{g}_j^{(t)}\right) \geq 0.5 \\ 1, & \text{elsewhere.} \end{cases} \tag{22}$$

Therefore, the MAP estimates of the channel wavelet coefficients are obtained by a simple denoising/thresholding rule as

$$g_j^{(t+1)} = \begin{cases} 0, & \text{if } \ \beta_j^{(t+1)} = 0 \\ \dfrac{\tau^2}{\alpha^2+\tau^2}\, \widetilde{g}_j^{(t+1)}, & \text{if } \ \beta_j^{(t+1)} = 1 \end{cases} \tag{23}$$

1) τ and λ updating: The prior parameters τ and λ stand respectively for the (significant)-wavelet coefficients energy and unsignificant coefficient probability. The update rules for these two parameters are MAP based rules derived from assigning conjugate priors to these parameters [10]:

$$\begin{aligned} \hat{\lambda} &= (\widetilde{L} - 1/2)/(L-1), \\ \hat{\tau}^2 &= \eta/(L-\widetilde{L}) \end{aligned} \tag{24}$$

where $\widetilde{L} = \mathrm{Card}\{j \mid \beta_j = 0\}$ and $\eta = \sum_{\beta_j=1} \left|g_j^{(t+1)}\right|^2$; $\mathrm{Card}\{.\}$ denoting the set cardinality.

2) Reduction of the number of estimated parameters: The thresholding procedure derived in this section, provides an easy framework for reducing the number of estimated coefficients. This can be done by discarding at each iteration, the elements of $\mathbf{g}^{(t+1)}$ that are replaced by zero in (23). The underlying assumption is as follows: whenever the estimator attributes an unknown wavelet coefficient to noise (replace it by zero), this coefficient will always be considered as noise and so will not be estimated in future iterations.

This operation is shown in figure 3 and can be modeled as:

$$\mathbf{g}_{\mathrm{tr}}^{(t+1)} = \Theta\big(\,\mathbf{g}^{(t+1)}\,\big), \quad \mathbf{T}_{\mathrm{tr}} = \Xi\big(\,\mathbf{T}\,\big) \tag{25}$$

where the truncation operator $\Theta(.)$ gathers in $\mathbf{g}_{\mathrm{tr}}^{(t+1)}$ the components of $\mathbf{g}^{(t+1)}$ that must be kept and the operator $\Xi(.)$ constructs $\mathbf{T}_{\mathrm{tr}}$ from $\mathbf{T}$ by keeping the rows corresponding to kept indexes. During the first iteration ($t = 0$), the algorithm does not perform any truncation and the EM algorithm estimates all coefficients. However, after each M-step, the number of unknown parameters to be estimated in the next iteration is reduced according (25) by using $\mathbf{g}_{\mathrm{tr}}^{(t+1)}$ and $\mathbf{T}_{\mathrm{tr}}$ in the update formula of the E-step (13).

IV. Decoding Method and Implementation Issues

According to equation (10), we make use of the information on transmitted symbols, obtained from the decoder, to update the channel estimate at each iteration. Besides, the decoder requires an estimate of the channel in order to provide the probability of encoded bits. Hence, the semi-blind channel estimation algorithm is naturally combined with the process of data decoding. The *a posteriori* probability of the unknown symbol S_k, $p(S_k|Y_k, \widehat{H}_k^{(t)})$, is calculated using the *a posteriori* probabilities provided by the decoder at the end of the t-th iteration as

$$p(S_k|Y_k, \widehat{H}) = \prod_{i=1}^{B} P_{\mathrm{dec}}(c_{k,i}) \tag{26}$$

where $P_{\mathrm{dec}}(c_{k,i})$ is the *a posteriori* probability corresponding to the i-th bit of S_k, $c_{k,i}$. At the first iteration, where no *a priori* information is available on bits $c_{k,i}$, $P_{\mathrm{dec}}(c_{k,i})$ are set to $1/2$.

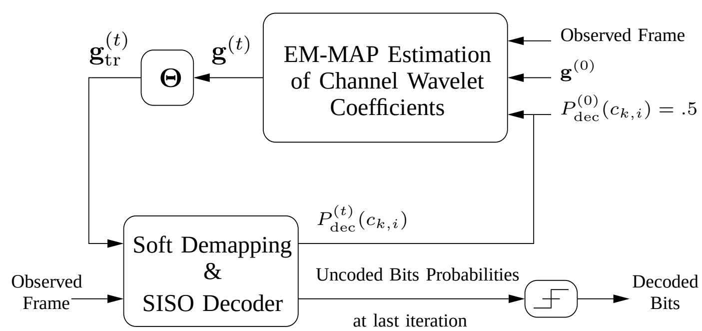

Fig. 3. EM-MAP channel estimation combined with the decoding process.

Among several possible ways to practically implement a joint channel estimation and decoding receiver, we adopt the following global procedure (see figure 3).

- Initialization ($t = 0$)
  - Set all probabilities of coded bits $P_{\mathrm{dec}}(c_{k,i})$ to $1/2$ and derive $p(\mathbf{S}|\mathbf{Y}, \mathbf{g}^{(0)})$ according to (26).
  - Initialize the unknown vector $\mathbf{g}$ by $\mathbf{g}^{(0)}$ obtained from pilot symbols.
- **for** $t = 1, \ldots, t_{\max}$
  - Use the current estimate $\mathbf{g}^{(t)}$ to calculate $\mathbf{g}^{(t+1)}$ according to (23).
  - Discard the wavelet coefficients that are replaced by zero for the next iteration by evaluating $\mathbf{g}_{\mathrm{tr}}^{(t)}$ and $\mathbf{T}_{\mathrm{tr}}$ from (25).
  - **if** $t \neq t_{\max}$: Use the current estimate $\mathbf{g}_{\mathrm{tr}}^{(t)}$ to update the probability of encoded bits $P_{\mathrm{dec}}(c_{k,i})$ and derive

$p(\mathbf{S}|\mathbf{Y}, \mathbf{g}^{(t)})$ from (26).
else: Decode the information data by thresholding the uncoded bit probabilities with 1/2.

V. SIMULATION RESULTS

In this section we present a comparative performance study of the proposed EM-MAP algorithm. The binary information data are encoded by a non-recursive non-systematic convolutional encoder with rate $R = 1/2$ and constraint length 3. Each frame has a payload of 1 KB along with 3 pilot symbols at the beginning for initializing the channel of each subband. The interleaver is random and operates over the entire frame. Among different wavelet families, "symmetric" wavelet basis functions [11] providing the sparser representation [9] have been considered.Unless otherwise mentioned, the curves are obtained after $t_{\max} = 4$ iterations.

First, a sparse channel model where only 20 wavelet coefficients out of total 96 have non zero values, is considered. The second channel, referred to as Corridor, is a line of sight (LOS) scenario issued from realistic UWB indoor channel measurements [12] where the receive and transmit antennas are located in a corridor separated by 9 meters.

Performance comparison is made with two pilot-only based approach using ML and minimum mean square error (MMSE) channel estimation, referred to as pilot-ML and pilot-MMSE. We also compare the proposed algorithm with two semi-blind channel estimation based on the EM algorithm, called respectively EM-Freq and EM-Wav. The first approach, consists of estimating the channel over all of the three subbands, using the model (3), similar to [5] while the second scheme is a wavelet domain EM based estimation of the channel where the prior model is set to have a uniform distribution.

Figure 4 depicts the mean square error (MSE) between true and estimated channel as a function of E_b/N_0. It can be noticed that, although the pilot-MMSE approach improves the estimation accuracy for low SNR values, the performance of pilot based channel estimation methods are very far from the family of semi-blind methods. Comparing the wavelet domain semi-blind approach (EM-Wav) and the frequency domain approach (EM-freq), shows that significant gain is achieved by the former method. As shown, the best performance is achieved by the EM-MAP method. We see that by using EM-MAP, a gain of almost 4 dB in SNR is achieved at MSE=2×10^{-3}, as compared to the EM-Wav method. This clearly shows the adequacy of the EM-MAP method for the case where the unknown channel has few non zero wavelet coefficients, which is in perfect agreement with the prior model.

Figure 5 shows the BER results along with the BER for the case of perfect channel state information (CSI). It can be seen that at a BER of 10^{-3}, the pilot-ML and the EM-Freq approaches are respectively 3.9 and 2 dB of SNR far from the BER obtained with the perfect channel. Furthermore, the performance of the Pilot-MMSE approach is not shown since it was very close to that of Pilot-ML. Also, we observe that wavelet based semi-blind methods perform closely to the perfect CSI case. For example, at BER=10^{-4}, the EM-MAP and EM-Wav method have respectively about 0.2 dB and 0.5 dB of SNR degradation from the performance obtained with perfect CSI.

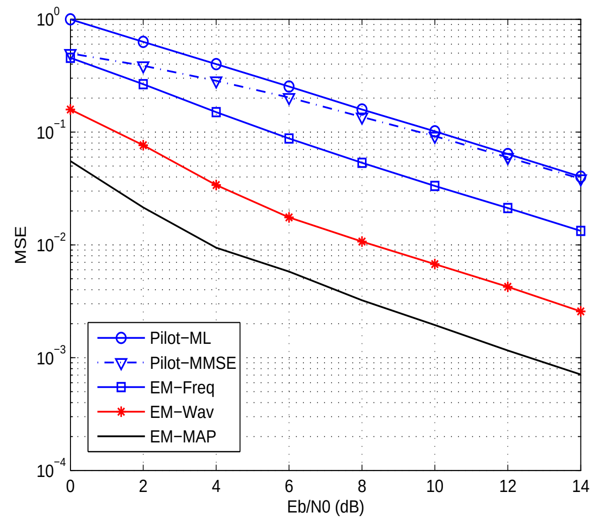

Fig. 4. Mean square error between the true and estimated coefficients for the sparse channel model.

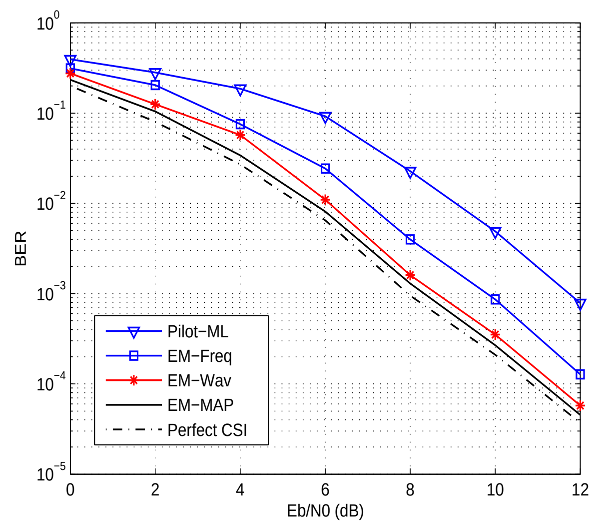

Fig. 5. BER performance of the EM-MAP method over the sparse channel model.

We now evaluate the performance of EM-MAP by considering the Corridor channel. Figure 6, shows that wavelet based methods again outperforms pilot based and EM-Freq methods in terms of MSE and BER. However, the EM-MAP performance is now comparable to that of EM-Wav method. This can be explained by noting that when the channel is not sparse, small values are attributed to λ by the algorithm (see (24)). This leads to a gaussian prior model with a large variance compared to the noise variance, which can be approximated with a uniform prior. As a results, the prior becomes less informative and the EM-MAP performs close

to EM-Wav, as shown in figures 6. Thus, the proposed EM-MAP algorithm is able to adapt its prior model parameters for each propagation environment.

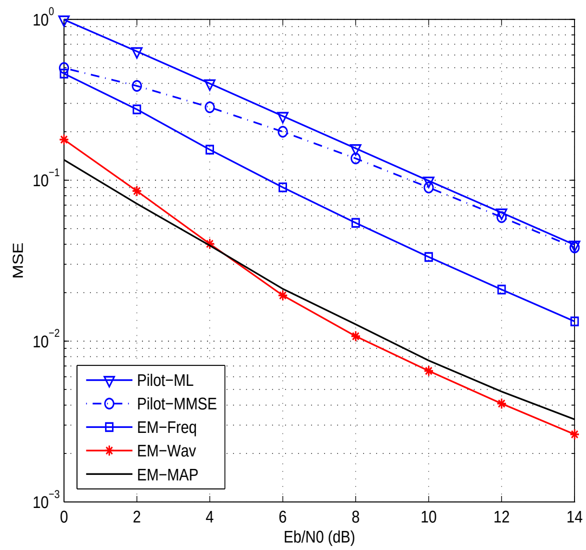

Fig. 6. Mean square error between the true and estimated coefficients over the Corridor channel.

Figure 7 shows the average number of estimated parameters versus the iteration number different channel scenarios. As observed, the EM-MAP approach tends to reduce significantly the number of estimated parameters. This can be seen for the sparse channel where the number of estimated parameters is reduced up to 20 parameters at the fifth iteration. Furthermore, under non-sparse Corridor channel, the figure shows that EM-MAP method is preferred to EM-Wav, due to its lower computational load.

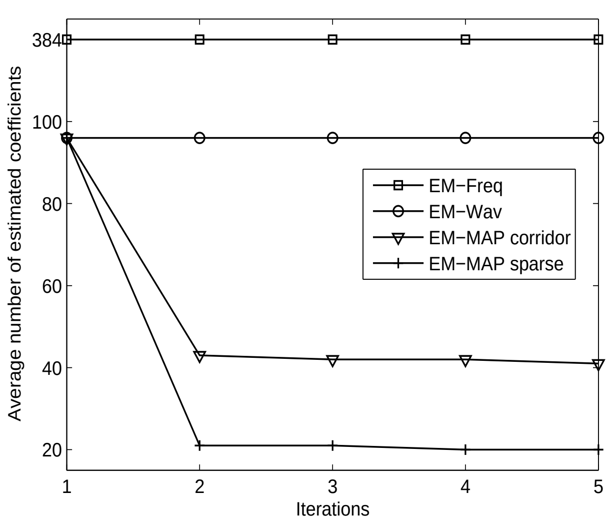

Fig. 7. Reduction of the number of estimated parameters through iterations, $Eb/N0 = 8$ dB.

VI. Conclusion

This paper proposed a semi-blind MAP channel estimation algorithm that integrates the advantages of wavelet based estimation. The investigated method naturally combines the EM iterations with the decoding process. We derived an equivalent data model for the multiband OFDM system involving the channel over all 3 subbands expressed in the wavelet domain. By choosing a Bernoulli-Gaussian prior distribution for the channel wavelet coefficients, the MAP estimator yields a thresholding procedure at the M-step of the EM algorithm which we used to reduce the number of estimated coefficients. With only few iterations, the EM-MAP method provides significant reduction in the number of estimated parameters and outperforms all considered pilot based and semi-blind methods.

References

[1] FCC, "First report and order, revision of part 15 of the commission's rules regarding ultra-wideband transmission systems," Tech. Rep., Feb. 2004.

[2] S. Roy, J. R. Foerster, V. S. Somayazulu, and D. G. Leeper, "Ultrawideband radio design: The promise of high-speed, short range wireless conectivity," *Proceedings of the IEEE*, vol. 92, February 2004.

[3] A. Batra, J. Balakrishnan, G. R. Aiello, J. R. Foerster, and A. Dabak, "Design of multiband OFDM system for realistic UWB channel environments," *IEEE transaction on microwave theory and techniques*, vol. 52, pp. 2123–2138, september 2004.

[4] R. Prasad, *OFDM for Wireless Communications Systems*, Artech House Publishers, september 2004.

[5] X. Ma, H. Kobayashi, and S. C. Schawrtz, "EM-based channel estimation algorithms for OFDM," *EURASIP Journal on Applied Signal Processing*, vol. 10, pp. 1460–1477, 2004.

[6] T. Y. Al-Naffouri, A. Bahai, and A. Paulraj, "Semi-blind channel identification and equalization in OFDM: An expectation-maximization approach," in *in Proc. IEEE Vehicular Technology Conference*, 2002, pp. 13–17.

[7] G. J. McLachlan and T. Krishnan, *The EM Algorithm and Extensions*, Wiley Series in probability and Statistics, Wiley, 1997.

[8] I. M. Johnstone and B. W. Silverman, "Neddles and straw in haystacks: Empirical bayes estimates of possibly sparse sequences," *Annals of Statistics*, vol. 32, pp. 1594–1649, 2004.

[9] S. Sadough and E. Jaffrot, "A wavelet packet based model for an ultra-wideband propagation channel," in *Proc. ECPS 2005*, Brest, France, March 15-18. 2005.

[10] J. N. Bernardo and A. F. M. Smith, *Bayesian Theory*, Wiley ans Sons, Chichester, England, 1994.

[11] S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, 1999.

[12] S. Bories, A. Sibille, and C. Roblin, "UWB indoor channel measurement," in *Proc. IWAT*, March 2005, pp. 466–469.